

Attorney Docket No. 00303.355US1

**DEAPROM AND TRANSISTOR WITH GALLIUM NITRIDE OR GALLIUM  
ALUMINUM NITRIDE GATE**

5

**Cross Reference To Related Applications**

This application is related to the following co-pending, commonly assigned U.S. patent applications: "DEAPROM HAVING AMORPHOUS SILICON CARBIDE GATE INSULATOR," serial number \_\_\_\_\_, "DYNAMIC ELECTRICALLY ALTERABLE PROGRAMMABLE READ ONLY MEMORY AND METHODS OF FABRICATION AND USE," serial number \_\_\_\_\_, "CARBURIZED SILICON GATE INSULATORS FOR INTEGRATED CIRCUITS," serial number \_\_\_\_\_, "SILICON CARBIDE GATE TRANSISTOR AND FABRICATION PROCESS," serial number \_\_\_\_\_, "TRANSISTOR WITH VARIABLE ELECTRON AFFINITY GATE AND METHODS OF FABRICATION AND USE," serial number \_\_\_\_\_, and "TRANSISTOR WITH SILICON OXYCARBIDE GATE AND METHODS OF FABRICATION AND USE," each of which is filed on even date herewith, and each of which disclosure is herein incorporated by reference.

20

**Field of the Invention**

The present invention relates generally to integrated circuit technology, including dynamic random access memories (DRAMs) and electrically erasable and programmable read only memories (EEPROMS), and particularly, but not by way of limitation, to a floating gate transistor memory that is dynamically electrically alterable and programmable.

**Background of the Invention**

Dynamic random access memories (DRAMs) are data storage devices that store data as charge on a storage capacitor. A DRAM typically includes an array of memory

cells. Each memory cell includes a storage capacitor and an access transistor for transferring charge to and from the storage capacitor. Each memory cell is addressed by a word line and accessed by a bit line. The word line controls the access transistor such that the access transistor controllably couples and decouples the storage capacitor to and from the bit line for writing and reading data to and from the memory cell.

The storage capacitor must have a capacitance that is large enough to retain a charge sufficient to withstand the effects of parasitic capacitances, noise due to circuit operation, and access transistor reverse-bias junction leakage currents between periodic data refreshes. Such effects can result in erroneous data. Obtaining a large capacitance 10 typically requires a storage capacitor having a large area. However, a major goal in DRAM design is to minimize the area of a DRAM memory cell to allow cells to be more densely packed on an integrated circuit die so that more data can be stored on smaller integrated circuits.

In achieving the goal of increasing DRAM array capacity by increasing cell 15 density, the sufficient capacitance levels of the DRAM storage capacitors must be maintained. A "stacked storage cell" design can increase the cell density to some degree. In this technique, two or more capacitor conductive plate layers, such as polycrystalline silicon (polysilicon or poly), are deposited over a memory cell access transistor on a semiconductor wafer. A high dielectric constant material is sandwiched 20 between these capacitor plate layers. Such a capacitor structure is known as a stacked capacitor cell (STC) because the storage capacitor plates are stacked on top of the access transistor. However, formation of stacked capacitors typically requires complicated process steps. Stacked capacitors also typically increase topographical features of the integrated circuit die, making subsequent lithography and processing, such as for 25 interconnection formation, more difficult. Alternatively, storage capacitors can be formed in deep trenches in the semiconductor substrate, but such trench storage capacitors also require additional process complexity. There is a need in the art to further increase memory storage density without adding process complexity or additional topography.

Electrically erasable and programmable read only memories (EEPROMs) provide nonvolatile data storage. EEPROM memory cells typically use field-effect transistors (FETs) having an electrically isolated (floating) gate that affects conduction between source and drain regions of the FET. A gate dielectric is interposed between the floating gate and an underlying channel region between source and drain regions. A control gate is provided adjacent to the floating gate, separated therefrom by an intergate dielectric.

In such memory cells, data is represented by charge stored on the polysilicon floating gates, such as by hot electron injection or Fowler-Nordheim tunneling during a write operation. Fowler-Nordheim tunneling is typically used to remove charge from the polysilicon floating gate during an erase operation. However, the relatively large electron affinity of the polysilicon floating gate presents a relatively large tunneling barrier energy at its interface with the underlying gate dielectric. The large tunneling barrier energy provides longer data retention times than realistically needed. For example, a data charge retention time at 85° C is estimated to be in millions of years for some floating gate memory devices. The large tunneling barrier energy also increases the voltages and time needed to store and remove charge to and from the polysilicon floating gate. "Flash" EEPROMs, which have an architecture that allows the simultaneous erasure of many floating gate transistor memory cells, require even longer erasure times to accomplish this simultaneous erasure. The large erasure voltages needed can result in hole injection into the gate dielectric. This can cause erratic overerasure, damage to the gate dielectric, and introduction of trapping states in the gate dielectric. The high electric fields that result from the large erasure voltages can also result in reliability problems, leading to device failure. There is a need in the art to obtain floating gate transistors that allow the use of lower programming and erasure voltages and shorter programming and erasure times.

Summary of the Invention

The present invention includes a memory cell that allows the use of lower programming and erasure voltages and shorter programming and erasure times by providing a gallium nitride (GaN) or gallium aluminum nitride (GaAlN) storage 5 electrode for storing charge. According to one aspect of the invention, the storage electrode provides a barrier energy, with an adjacent insulator, of less than approximately 3.3 eV, which provides improved erase times or voltages.

In one embodiment, the memory cell includes a floating gate transistor, having a GaN or GaAlN floating gate. A refresh circuit allows dynamic refreshing of charge 10 stored on the floating gate. The barrier energy can be lowered to a desired value by selecting the appropriate material composition of the floating gate. As a result, lower programming and erasure voltages and shorter programming and erasure times are obtained.

Another aspect of the present invention provides a method of using a floating 15 gate transistor having a reduced barrier energy between a floating gate electrode and an adjacent insulator. Data is stored by changing the charge of the floating gate. Data is refreshed based on a data charge retention time established by the barrier energy between the GaN or GaAlN floating gate and the adjacent insulator. Data is read by detecting a conductance between a source and a drain. The large transconductance gain 20 of the memory cell of the present invention provides a more easily detected signal and reduces the required data storage capacitance value and memory cell size when compared to a conventional dynamic random access memory (DRAM) cell.

The present invention also includes a method of forming a floating gate 25 transistor. Source and drain regions are formed. A gate insulator is formed. A GaN or GaAlN floating gate is formed, such that the floating gate is isolated from conductors and semiconductors. In one embodiment, the floating gate is formed by metal organic chemical vapor deposition (MOCVD). In another embodiment, the floating gate is formed by plasma-enhanced molecular beam epitaxy (PEMBE). The GaN or GaAlN floating gate provides a relatively short data charge retention time, but advantageously

provides a shorter write/programming and erase times, making operation of the present memory speed competitive with a DRAM.

The present invention also includes a memory device that is capable of providing short programming and erase times, low programming and erase voltages, 5 and lower electric fields in the memory cell for improved reliability. The memory device includes a plurality of memory cells. Each memory cell includes a transistor. Each transistor includes a source region, a drain region, a channel region between the source and drain regions, and a GaN or GaAlN floating gate that is separated from the channel region by an insulator. The transistor also includes a control gate located 10 adjacent to the floating gate and separated therefrom by an intergate dielectric. According to one aspect of the invention, a refresh circuit is provided. The memory device includes flash electrically erasable and programmable read only memory (EEPROM), dynamic random access memory (DRAM), and dynamically electrically alterable and programmable read only memory (DEAPROM) embodiments.

15 The memory cell of the present invention provides a reduced barrier energy, large transconductance gain, an easily detected signal, and reduces the required data storage capacitance value and memory cell size. The lower barrier energy increases tunneling current and also advantageously reduces the voltage required for writing and erasing the floating gate transistor memory cells. For example, conventional polysilicon 20 floating gate transistors typically require complicated and noisy on-chip charge pump circuits to generate the large erasure voltage, which typically far exceeds other voltages required on the integrated circuit. The present invention allows the use of lower erasure voltages that are more easily provided by simpler on-chip circuits. Reducing the erasure voltage also lowers the electric fields, minimizing reliability problems that can lead to 25 device failure, and better accommodating downward scaling of device dimensions. Alternatively, the thickness of the gate insulator can be increased from the typical thickness of a silicon dioxide gate insulator to improve reliability or simplify processing, since the lower barrier energy allows easier transport of charge across the gate insulator by Fowler-Nordheim tunneling.

According to another aspect of the invention, the shorter retention time of data charges on the GaN or GaAlN floating electrode, resulting from the smaller barrier energy, is accommodated by refreshing the data charges on the floating electrode. By decreasing the data charge retention time and periodically refreshing the data, the write and erase operations can be several orders of magnitude faster such that the present memory is speed competitive with a DRAM. In this respect, the memory operates similar to a memory cell in DRAM, but avoids the process complexity, additional space needed, and other limitations of forming stacked or trench DRAM capacitors.

The memory cell of the present invention can be made smaller than a conventional DRAM memory cell. Moreover, because the storage capacitor of the present invention is integrally formed as part of the transistor, rather than requiring complex and costly non-CMOS stacked and trench capacitor process steps, the memory of the present invention should be cheaper to fabricate than DRAM memory cells, and should more easily scale downward as CMOS technology advances.

15

#### Brief Description of the Drawings

In the drawings, like numerals describe substantially similar components throughout the several views.

Figure 1 is a simplified schematic/block diagram illustrating generally one embodiment of a memory including reduced barrier energy floating electrode memory cells.

Figure 2 is a cross-sectional view that illustrates generally a floating gate transistor embodiment of a memory cell provided by the present invention.

Figure 3 is an energy band diagram that illustrates generally conduction band energy levels in a floating gate transistor provided by the present invention.

Figure 4 is a graph comparing barrier energy vs. tunneling distance for a conventional floating gate transistor and one embodiment of a the present invention having a lower barrier energy.

Figure 5 is a graph that illustrates generally the relationship between Fowler-Nordheim tunneling current density vs. the barrier energy  $\Phi_{Gr}$  at various parameterized values  $E_1 < E_2 < E_3$  of an electric field.

Figure 6 illustrates generally how the barrier energy affects the time needed to 5 perform write and erase operations by Fowler-Nordheim tunneling for a particular voltage.

Figure 7 is a graph that illustrates generally charge density vs. write/erase time for three different embodiments of a floating gate FET.

Figure 8 is a cross-sectional view, similar to Figure 2, but having a larger area 10 control gate - floating gate capacitor than the floating gate - substrate capacitor.

Figure 9A is a schematic diagram, labeled prior art, that illustrates generally a conventional DRAM memory cell.

Figure 9B is a schematic diagram that illustrates generally one embodiment of a floating gate FET memory cell according to the present invention.

15

#### Detailed Description of the Invention

In the following detailed description of the invention, reference is made to the accompanying drawings which form a part hereof, and in which is shown, by way of illustration, specific embodiments in which the invention may be practiced. In the 20 drawings, like numerals describe substantially similar components throughout the several views. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention. Other embodiments may be utilized and structural, logical, and electrical changes may be made without departing from the scope of the present invention. The terms wafer and substrate used in the following 25 description include any semiconductor-based structure having an exposed surface with which to form the integrated circuit structure of the invention. Wafer and substrate are used interchangeably to refer to semiconductor structures during processing, and may include other layers that have been fabricated thereupon. Both wafer and substrate include doped and undoped semiconductors, epitaxial semiconductor layers supported

by a base semiconductor or insulator, as well as other semiconductor structures well known to one skilled in the art. The following detailed description is, therefore, not to be taken in a limiting sense, and the scope of the present invention is defined only by the appended claims.

5        The present invention discloses a dynamic electrically alterable programmable read only memory (DEAPROM) cell. The memory cell has a floating electrode, which is defined as an electrode that is "electrically isolated" from conductors and semiconductors by an insulator such that charge storage upon and removal from the floating electrode depends upon charge conduction through the insulator. In one 10 embodiment, described below, the floating electrode is a floating gate electrode in a floating gate field-effect transistor, such as used in flash electrically erasable and programmable read only memories (EEPROMs). However, a capacitor or any other structure having a floating electrode and adjacent insulator could also be used according to the techniques of the present invention described below. According to one aspect of 15 the present invention, a barrier energy between the floating electrode and the insulator is lower than the barrier energy between polycrystalline silicon (polysilicon) and silicon dioxide ( $\text{SiO}_2$ ), which is approximately 3.3 eV. According to another aspect of the present invention, the shorter retention time of data charges on the floating electrode, resulting from the smaller barrier energy, is accommodated by refreshing the data 20 charges on the floating electrode. In this respect, the memory operates similar to a memory cell in a dynamic random access memory (DRAM). These and other aspects of the present invention are described in more detail below.

Figure 1 is a simplified schematic/block diagram illustrating generally one embodiment of a memory 100 according to one aspect of the present invention, in 25 which reduced barrier energy floating electrode memory cells are incorporated. Memory 100 is referred to as a dynamic electrically alterable programmable read only memory (DEAPROM) in this application, but it is understood that memory 100 possesses certain characteristics that are similar to DRAMs and flash EEPROMs, as explained below. For a general description of how a flash EEPROM operates, see B.

Dipert et al., "Flash Memory Goes Mainstream," IEEE Spectrum, pp. 48-52 (Oct. 1993), which is incorporated herein by reference. Memory 100 includes a memory array 105 of multiple memory cells 110. Row decoder 115 and column decoder 120 decode addresses provided on address lines 125 to access the addressed memory cells in 5 memory array 105. Command and control circuitry 130 controls the operation of memory 100 in response to control signals received on control lines 135 from a processor 140 or other memory controller during read, write, refresh, and erase operations. Command and control circuitry 130 includes a refresh circuit for periodically refreshing the data stored on floating gate transistor or other floating 10 electrode memory cells 110. Voltage control 150 provides appropriate voltages to the memory cells during read, write, refresh, and erase operations. Memory 100, as illustrated in Figure 1, has been simplified for the purpose of illustrating the present invention and is not intended to be a complete description. Only the substantial differences between DEAPROM memory 100 and conventional DRAM and flash 15 EEPROM memories are discussed below.

Figure 2 is a cross-sectional view that illustrates generally, by way of example, but not by way of limitation, one floating gate transistor embodiment of a memory cell 110. Other structural arrangements of floating gate transistors are included within the present invention. Also included are any memory cells that incorporate a floating 20 electrode (such as a floating electrode capacitor) having, at an interface between the floating electrode an adjacent insulator, a barrier energy that is less than the barrier energy at a polysilicon-SiO<sub>2</sub> interface. In the embodiment of Figure 2, memory cell 110 includes a floating gate FET 200, which is illustrated as an n-channel FET, but understood to include a p-channel FET embodiment as well.

25 FET 200 includes a source 205, a drain 210, a floating gate 215 electrode, and a control gate 220 electrode. A gate insulator 225 is interposed between floating gate 215 and substrate 230. An intergate insulator 235 is interposed between floating gate 215 and control gate 220. In one embodiment, substrate 230 is a bulk semiconductor, such as silicon. In another embodiment, substrate 230 includes a thin semiconductor surface

layer formed on an underlying insulating portion, such as in a semiconductor-on-insulator (SOI) or other thin film transistor technology. Source 205 and drain 210 are formed by conventional complementary metal-oxide-semiconductor (CMOS) processing techniques. Source 205 and drain 210 are separated by a predetermined 5 length for forming an inversion channel 240 therebetween.

Figure 3 is an energy band diagram that illustrates generally the conduction band energy levels in floating gate 215, gate insulator 225, and substrate 230. Electron affinities  $\chi_{215}$ ,  $\chi_{225}$ , and  $\chi_{230}$  describe floating gate 215, gate insulator 225, and substrate 230, respectively, when measured with respect to a vacuum level 300. A barrier energy 10  $\Phi_{GI}$ , which describes the barrier energy at the interface between floating gate 215 and gate insulator 225, is given by a difference in electron affinities, as illustrated in Equation 1.

$$\Phi_{GI} = \chi_{215} - \chi_{225} \quad (1)$$

A barrier energy  $\Phi_{SG}$ , which describes the barrier energy at the interface between substrate 230 and gate insulator 225, is given by a difference in electron affinities, as 15 illustrated in Equation 2.

$$\Phi_{SG} = \chi_{230} - \chi_{225} \quad (2)$$

Silicon (monocrystalline or polycrystalline Si) has an electron affinity  $\chi_{215} \approx 4.2$  eV. Silicon dioxide ( $\text{SiO}_2$ ) has an electron affinity,  $\chi_{225}$ , of about 0.9 eV. The resulting barrier energy at a conventional Si- $\text{SiO}_2$  interface between a floating gate and a gate insulator is approximately equal to 3.3 eV. One aspect of the present invention provides 20 a barrier energy  $\Phi_{GI}$  that is less than the 3.3 eV barrier energy of a conventional Si- $\text{SiO}_2$  interface.

According to one aspect of the invention, the interface between floating gate 215 and gate insulator 225 provides a smaller barrier energy  $\Phi_{GI}$  than the 3.3 eV barrier energy at an interface between polysilicon and silicon dioxide, such as by an appropriate

selection of the material composition of one or both of floating gate 215 and gate insulator 225. In one embodiment, the smaller barrier energy  $\Phi_{GI}$  is obtained by forming floating gate 215 from a material having a smaller electron affinity  $\chi_{215}$  than polysilicon. In one embodiment, for example, polycrystalline or microcrystalline 5 silicon carbide (SiC) is used as the material for forming floating gate 215. In another embodiment, the smaller barrier energy  $\Phi_{GI}$  is obtained by forming gate insulator 225 from a material having a higher electron affinity  $\chi_{225}$  than  $\text{SiO}_2$ . In one embodiment, for example, amorphous SiC is used as the material for forming gate insulator 225. In yet another embodiment, the smaller barrier energy  $\Phi_{GI}$  is obtained by a combination of 10 forming floating gate 215 from a material having a smaller electron affinity  $\chi_{215}$  than polysilicon and also forming gate insulator 225 from a material having a higher electron affinity  $\chi_{225}$  than  $\text{SiO}_2$ .

The smaller barrier energy  $\Phi_{GI}$  provides current conduction across gate insulator 225 that is easier than for a polysilicon- $\text{SiO}_2$  interface. The present invention includes 15 any mechanism of providing such easier current conduction across gate insulator 225, including, but not limited to "hot" electron injection, thermionic emission, Schottky emission, Frenkel-Poole emission, and Fowler-Nordheim tunneling. Such techniques for transporting charge carriers across an insulator, such as gate insulator 225, are all enhanced by providing a smaller barrier energy  $\Phi_{GI}$  according to the techniques of the 20 present invention. These techniques allow increased current conduction, current conduction at lower voltages across gate insulator 225 and lower electric fields in gate insulator 225, shorter data write and erase times, use of a thicker and more reliable gate insulator 225, and other advantages explained below.

Figure 4 is a graph illustrating generally barrier energy versus tunneling distance 25 for a conventional polysilicon- $\text{SiO}_2$  interface having a 3.3 eV barrier energy. Figure 4 also illustrates barrier energy versus tunneling distance for an interface according to the present invention that has a barrier energy of  $\Phi_{GI} \approx 1.08$  eV, which is selected as an illustrative example, and not by way of limitation. The smaller barrier energy  $\Phi_{GI}$  reduces the energy to which the electrons must be excited to be stored on or removed

from the floating gate 215, such as by thermal emission over the barrier. The smaller barrier energy  $\Phi_{GI}$  also reduces the distance that electrons have to traverse, such as by Fowler-Nordheim tunneling, to be stored upon or removed from floating gate 215. In Figure 4, "do" represents the tunneling distance of a conventional floating gate 5 transistor due to the 3.3 eV barrier energy represented by the dashed line "OLD". The tunneling distance "dn" corresponds to a floating gate transistor according to the present invention and its smaller barrier energy, such as  $\Phi_{GI} \approx 1.08$  eV, for example, represented by the dashed line "NEW". Even a small reduction in the tunneling distance results in a 10 large increase in the tunneling probability, as described below, because the tunneling probability is an exponential function of the reciprocal of the tunneling distance.

The Fowler-Nordheim tunneling current density in gate insulator 225, which is illustrated approximately by Equation 3 below, is described in a textbook by S.M. Sze, "Physics of Semiconductor Devices," John Wiley & Sons, New York (1969), p. 496.

$$J = AE^2 e^{(-\frac{B}{E})} \quad (3)$$

In Equation 3, J is the current density in units of amperes/cm<sup>2</sup>, E is the electric field in 15 gate insulator 225 in units of volts/cm and A and B are constants, which are particular to the material of gate insulator 225, that depend on the effective electron mass in the gate insulator 225 material and on the barrier energy  $\Phi_{GI}$ . The constants A and B scale with the barrier energy  $\Phi_{GI}$ , as illustrated approximately by Equations 4 and 5, which are disclosed in S.R. Pollack et al., "Electron Transport Through Insulating Thin Films," 20 Applied Solid State Science, Vol. 1, Academic Press, New York, (1969), p. 354.

$$A \propto \left( \frac{1}{\Phi_{GI}} \right) \quad (4)$$

$$B\alpha(\Phi_{GI})^{\frac{3}{2}} \quad (5)$$

For a conventional floating gate FET having a 3.3 eV barrier energy at the interface between the polysilicon floating gate and the  $\text{SiO}_2$  gate insulator,  $A = 5.5 \times 10^{-16}$  amperes/Volt<sup>2</sup> and  $B = 7.07 \times 10^7$  Volts/cm, as disclosed in D.A. Baglee, "Characteristics and Reliability of 100 Å Oxides," Proc. 22nd Reliability Symposium, (1984), p. 152. One aspect of the present invention includes selecting a smaller barrier energy  $\Phi_{GI}$  such as, by way of example, but not by way of limitation,  $\Phi_{GI} \approx 1.08$  eV. The constants  $A$  and  $B$  for  $\Phi_{GI} \approx 1.08$  eV can be extrapolated from the constants  $A$  and  $B$  for the 3.3 eV polysilicon- $\text{SiO}_2$  barrier energy using Equations 4 and 5. The barrier energy  $\Phi_{GI} \approx 1.08$  eV yields the resulting constants  $A = 1.76 \times 10^{-15}$  amperes/Volt<sup>2</sup> and  $B = 1.24 \times 10^7$  Volts/cm.

Figure 5 is a graph that illustrates generally the relationship between Fowler-Nordheim tunneling current density vs. the barrier energy  $\Phi_{GI}$ , such as at various parameterized values  $E_1 < E_2 < E_3$  of an electric field in gate insulator 225. The tunneling current density increases as electric field is increased. The tunneling current also increases by orders of magnitude as the barrier energy  $\Phi_{GI}$  is decreased, such as by selecting the materials for floating gate 215 and gate insulator 225 or otherwise reducing the barrier energy  $\Phi_{GI}$  according to the techniques of the present invention. In particular, Figure 5 illustrates a comparison between tunneling current densities at the 3.3 eV barrier energy of a conventional polysilicon- $\text{SiO}_2$  interface and at the illustrative example barrier energy  $\Phi_{GI} \approx 1.08$  eV for which constants  $A$  and  $B$  were extrapolated above. Reducing the 3.3 eV barrier energy to  $\Phi_{GI} \approx 1.08$  eV increases the tunneling current density by several orders of magnitude.

Figure 6 is a conceptual diagram, using rough order of magnitude estimates, that illustrates generally how the barrier energy affects the time needed to perform write and erase operations by Fowler-Nordheim tunneling for a particular voltage, such as across gate insulator 225. Figure 6 also illustrates how the barrier energy affects data charge 5 retention time, such as on floating gate 215 at a temperature of 250 degrees Celsius. Both write and erase time 600 and data charge retention time 605 are decreased by orders of magnitude as the barrier energy is decreased, according to the present invention, from the conventional polysilicon-SiO<sub>2</sub> interface barrier energy of 3.3 eV to the illustrative example lower barrier energy  $\Phi_{GI} \approx 1.08$  eV for which constants A and B 10 were extrapolated above.

The lower barrier energy  $\Phi_{GI}$  and increased tunneling current advantageously provides faster write and erase times. This is particularly advantageous for "flash" EEPROMs or DEAPROMs in which many floating gate transistor memory cells must be erased simultaneously, requiring a longer time to transport the larger quantity of 15 charge. For a flash EEPROM using a polysilicon floating gate transistor having an underlying SiO<sub>2</sub> gate insulator 225, the simultaneous erasure of a block of memory cells requires a time that is on the order of milliseconds. The write and erase time of the floating gate FET 200 is illustrated approximately by Equation 6.

$$t = \int_0^t dt = \int_0^Q \left( \frac{1}{J_{225} - J_{235}} \right) dQ \quad (6)$$

In Equation 6, t is the write/erase time,  $J_{225}$  and  $J_{235}$  are the respective tunneling current 20 densities in gate dielectric 225 and intergate dielectric 235, Q is the charge density in Coulombs/cm<sup>2</sup> on floating gate 215. Equation 6 is evaluated for a specific voltage on control gate 220 using Equations 7 and 8.

$$E_{225} = \frac{V_{220}}{\left[ d_{225} + d_{235} \left( \frac{\epsilon_{225}}{\epsilon_{235}} \right) \right] - \frac{Q}{\left[ \epsilon_{225} + \epsilon_{235} \left( \frac{d_{225}}{d_{235}} \right) \right]}} \quad (7)$$

$$E_{235} = \frac{V_{220}}{\left[ d_{235} + d_{225} \left( \frac{\epsilon_{235}}{\epsilon_{225}} \right) \right] + \frac{Q}{\left[ \epsilon_{235} + \epsilon_{225} \left( \frac{d_{235}}{d_{225}} \right) \right]}} \quad (8)$$

In Equations 7 and 8,  $V_{220}$  is the voltage on control gate 220,  $E_{225}$  and  $E_{235}$  are the respective electric fields in gate insulator 225 and intergate insulator 235,  $d_{225}$  and  $d_{235}$  are the respective thicknesses of gate insulator 225 and intergate insulator 235, and  $\epsilon_{225}$  and  $\epsilon_{235}$  are the respective permittivities of gate insulator 225 and intergate insulator 235.

Figure 7 is a graph that illustrates generally charge density vs. write/erase time for three different embodiments of the floating gate FET 200, each of which have a polysilicon floating gate 215, by way of illustrative example. Line 700 illustrates generally, by way of example, but not by way of limitation, the charge density vs. write/erase time obtained for a floating gate FET 200 having a 100 Å SiO<sub>2</sub> gate insulator 225 and a 150 Å SiO<sub>2</sub> (or thinner oxynitride equivalent capacitance) intergate insulator 235.

Line 705 is similar to line 700 in all respects except that line 705 illustrates a floating gate FET 200 in which gate insulator 225 comprises a material having a higher electron affinity  $\chi_{225}$  than SiO<sub>2</sub>, thereby providing a lower barrier energy  $\Phi_{Gr}$  at the

interface between polysilicon floating gate 215 and gate insulator 225. The increased tunneling current results in shorter write/erase times than those illustrated by line 700.

Line 710 is similar to line 705 in all respects except that line 710 illustrates a floating gate FET 200 in which gate insulator 225 has a lower barrier energy  $\Phi_{GI}$  than 5 for line 705, or intergate insulator 235 has a higher permittivity  $\epsilon_{235}$  than for line 705, or control gate 220 has a larger area than floating gate 215, such as illustrated by way of example by the floating gate FET 800 in the cross-sectional view of Figure 8. As seen in Figure 8, the area of a capacitor formed by the control gate 220, the floating gate 215, and the intergate insulator 235 is larger than the area of a capacitor formed by the 10 floating gate 215, the gate insulator 225, and the inversion channel 240 underlying gate insulator 225. Alternatively, or in combination with the techniques illustrated in Figure 8, the intergate insulator 235 can have a higher permittivity than the permittivity of silicon dioxide.

As illustrated in Figure 7, the barrier energy  $\Phi_{GI}$  can be selected to reduce the 15 write/erase time. In one embodiment, by way of example, but not by way of limitation, the barrier energy  $\Phi_{GI}$  is selected to obtain a write/erase time of less than or equal to 1 second, as illustrated in Figure 7. In another embodiment, by way of example, but not by way of limitation, the barrier energy  $\Phi_{GI}$  is selected to obtain a write/erase time of less than or equal to 1 millisecond, as illustrated in Figure 7. Other values of write/erase 20 time can also be obtained by selecting the appropriate value of the barrier energy  $\Phi_{GI}$ .

The lower barrier energy  $\Phi_{GI}$  and increased tunneling current also advantageously reduces the voltage required for writing and erasing the floating gate transistor memory cells 110. For example, conventional polysilicon floating gate transistors typically require complicated and noisy on-chip charge pump circuits to 25 generate the large erasure voltage, which typically far exceeds other voltages required on the integrated circuit. The present invention allows the use of lower erasure voltages that are more easily provided by simpler on-chip circuits. Reducing the erasure voltage also lowers the electric fields, minimizing reliability problems that can lead to device failure, and better accommodating downward scaling of device dimensions. In one

embodiment, the barrier energy  $\Phi_{\text{Gr}}$  is selected, as described above, to obtain an erase voltage of less than the 12 Volts required by typical EEPROM memory cells.

Alternatively, the thickness of the gate insulator 225 can be increased from the typical thickness of a silicon dioxide gate insulator to improve reliability or simplify 5 processing, since the lower barrier energy  $\Phi_{\text{Gr}}$  allows easier transport of charge across the gate insulator 225 by Fowler-Nordheim tunneling.

The lower barrier energy  $\Phi_{\text{Gr}}$  also decreases the data charge retention time of the charge stored on the floating gate 215, such as from increased thermal excitation of stored charge over the lower barrier  $\Phi_{\text{Gr}}$ . However, conventional polysilicon floating 10 gates and adjacent  $\text{SiO}_2$  insulators (e.g., 90 Å thick) have a data charge retention time estimated in the millions of years at a temperature of 85 degrees C, and estimated in the 1000 hour range even at extremely high temperatures such as 250 degrees C. Since such long data charge retention times are longer than what is realistically needed, a shorter data charge retention time can be accommodated in order to obtain the benefits 15 of the smaller barrier energy  $\Phi_{\text{Gr}}$ . In one embodiment of the present invention, by way of example, but not by way of limitation, the barrier energy  $\Phi_{\text{Gr}}$  is lowered to  $\Phi_{\text{Gr}} \approx 1.08$  eV by appropriately selecting the composition of the materials of floating gate 215 and gate insulator 225, as described below. As a result, an estimated data charge retention 20 time of approximately 40 seconds at a high temperature, such as 250 degrees C, is obtained.

According to one aspect of the present invention, the data stored on the DEAPROM floating gate memory cell 110 is periodically refreshed at an interval that is shorter than the data charge retention time. In one embodiment, for example, the data is refreshed every few seconds, such as for an embodiment having a high temperature 25 retention time of approximately 40 seconds for  $\Phi_{\text{Gr}} \approx 1.08$  eV. The exact refresh rate can be experimentally determined and tailored to a particular process of fabricating the DEAPROM. By decreasing the data charge retention time and periodically refreshing the data, the write and erase operations can be several orders of magnitude faster, as described above with respect to Figure 7.

Figures 9A and 9B are schematic diagrams that respectively illustrate generally a conventional DRAM memory cell and the present invention's floating gate FET 200 embodiment of memory cell 110. In Figure 9A, the DRAM memory cell includes an access FET 900 and stacked or trench storage capacitor 905. Data is stored as charge on storage capacitor 905 by providing a control voltage on control line 910 to activate FET 900 for conducting charge. Data line 915 provides a write voltage to conduct charge across FET 900 for storage on storage capacitor 905. Data is read by providing a control voltage on control line 910 to activate FET 900 for conducting charge from storage capacitor 905, thereby incrementally changing a preinitialized voltage on data line 915. The resulting small change in voltage on data line 915 must be amplified by a sense amplifier for detection. Thus, the DRAM memory cell of Figure 9A inherently provides only a small data signal. The small data signal is difficult to detect.

In Figure 9B, the DEAPROM memory cell 110 according to the present invention includes floating gate FET 200, having source 205 coupled to a ground voltage or other reference potential. Data is stored as charge on floating gate 215 by providing a control voltage on control line 920 and a write voltage on data line 925 for hot electron injection or Fowler-Nordheim tunneling. This is similar to conventional EEPROM techniques, but advantageously uses the reduced voltages and/or a shorter write time of the present invention.

The DEAPROM memory cell 110 can be smaller than the DRAM memory cell of Figure 9A, allowing higher density data storage. The leakage of charge from floating gate 215 can be made less than the reverse-bias junction leakage from storage capacitor 905 of the DRAM memory cell by tailoring the barrier energy  $\Phi_{Gr}$  according to the techniques of the present invention. Also, the DEAPROM memory cell advantageously uses the large transconductance gain of the floating gate FET 200. The conventional DRAM memory cell of Figure 9A provides no such gain; it is read by directly transferring the data charge from storage capacitor 905. By contrast, the DEAPROM memory cell 110 is read by placing a read voltage on control line 920, and detecting the current conducted through FET 200, such as at data line 925. The current conducted

through FET 200 changes significantly in the presence or absence of charge stored on floating gate 215. Thus, the present invention advantageously provides an large data signal that is easy to detect, unlike the small data signal provided by the conventional DRAM memory cell of Figure 9A.

5 For example, the current for floating gate FET 200 operating in the saturation region can be approximated by Equation 9.

$$I_{DS} = \frac{1}{2} \mu C_o \left( \frac{W}{L} \right) (V_G - V_T)^2 \quad (9)$$

In Equation 9,  $I_{DS}$  is the current between drain 210 and source 205,  $C_o$  is the capacitance per unit area of the gate insulator 225,  $W/L$  is the width/length aspect ratio of FET 200,  $V_G$  is the gate voltage applied to control gate 220, and  $V_T$  is the turn-on threshold 10 voltage of FET 200.

For an illustrative example, but not by way of limitation, a minimum-sized FET having  $W/L=1$ , can yield a transconductance gain of approximately  $71 \mu\text{A/Volt}$  for a typical process. In this illustrative example, sufficient charge is stored on floating gate 215 to change the effective threshold voltage  $V_T$  by approximately 1.4 Volts, thereby 15 changing the current  $I_{DS}$  by approximately 100 microamperes. This significant change in current can easily be detected, such as by sampling or integrating over a time period of approximately 10 nanoseconds, for example, to obtain a detected data charge signal of 1000 fC. Thus, the DEAPROM memory cell 110 is capable of yielding a detected 20 data charge signal that is approximately an order of magnitude larger than the typical 30 fC to 100 fC data charges typically stored on DRAM stacked or trench capacitors.

Since DEAPROM memory cell 110 requires a smaller capacitance value than a conventional DRAM memory cell, DEAPROM memory cell 110 can be made smaller than a conventional DRAM memory cell. Moreover, because the CMOS-compatible DEAPROM storage capacitor is integrally formed as part of the transistor, rather than 25 requiring complex and costly non-CMOS stacked and trench capacitor process steps, the

DEAPROM memory of the present invention should be cheaper to fabricate than DRAM memory cells, and should more easily scale downward as CMOS technology advances.

5

### GaN and GaAlN Gate Material Embodiments

In one embodiment, the present invention provides a DEAPROM having a memory cell 110 including a FET 200 having an at least partially crystalline (e.g., monocrystalline, polycrystalline, microcrystalline, nanocrystalline, or combination thereof) gallium nitride (GaN) or gallium aluminum nitride (GaAlN) floating gate 215.

10 In one embodiment, a composition  $v$  of a polycrystalline  $\text{Ga}_{1-v}\text{Al}_v\text{N}$  floating gate 215 is selected approximately between  $0 < v < 1$  to obtain a desired barrier energy, as described below. The GaAlN floating gate 215 provides a lower electron affinity than polysilicon. The GaAlN floating gate 215 electron affinity can be approximately between  $0.6 \text{ eV} < \chi_{215} < 2.7 \text{ eV}$  as the GaAlN composition variable  $v$  is decreased from 15 1 to 0. See V. M. Bermudez et al. "The Growth and Properties of Al and AlN films on GaN" *J. Appl. Physics*, Vol. 79, No. 1, pp. 110-119 (1996). As a result, the GaAlN floating gate 215 provides a smaller resulting barrier energy  $\Phi_{\text{Gr}}$  than a polysilicon gate material having an electron affinity  $\chi_{215} \approx 4.2 \text{ eV}$ . For example, using a  $\text{SiO}_2$  gate insulator 225, a barrier energy approximately between  $-0.3 \text{ eV} < \Phi_{\text{Gr}} < 1.8 \text{ eV}$  is 20 obtained using an GaAlN floating gate 215 as the GaAlN composition  $v$  varies between  $v \approx 1$  (i.e., approximately AlN) and  $v \approx 0$  (i.e., approximately GaN). By contrast, a conventional polysilicon floating gate material provides a barrier energy  $\Phi_{\text{Gr}} \approx 3.3 \text{ eV}$  at an interface with an  $\text{SiO}_2$  gate insulator 225.

25 In one embodiment, substrate 230 is bulk silicon, although other bulk semiconductor and semiconductor-on-insulator (SOI) materials could also be used for substrate 230 such as, for example, sapphire, gallium arsenide (GaAs), GaN, AlN, and diamond. In one embodiment, gate insulator 225 is  $\text{SiO}_2$ , although other dielectric materials could also be used for gate insulator 225, as described above, such as amorphous insulating GaN (a-GaN), and amorphous insulating AlN (a-AlN). The FET

200 using a GaAlN floating gate 215 has mobility and turn-on threshold voltage ( $V_T$ ) magnitude parameters that are advantageously influenced less by charge at  $\text{SiO}_2$ -GaAlN interface surface states than at a conventional  $\text{SiO}_2$ -polysilicon interface.

In one embodiment floating gate 215 is formed of a polycrystalline, 5 microcrystalline, or nanocrystalline, GaN thin film that is CVD deposited on a thin (e.g., 500 Å thick) AlN buffer layer, such as by metal organic chemical vapor deposition (MOCVD), which advantageously yields improved crystal quality and reduced microscopic fluctuation of crystallite orientation. See e.g., V. M. Bermudez et al. "The Growth and Properties of Al and AlN films on GaN" J. Appl. Physics, Vol. 79, No. 1, 10 pp. 110-119 (1996). See also I. Akasaki et al. "Effects of AlN Buffer Layer on Crystallographic Structure and On Electrical and Optical Properties of GaN and  $\text{Ga}_1$ <sub>x</sub> $\text{Al}_x\text{N}$  Films Grown on Sapphire Substrate by MOVPE," J. Of Crystal Growth, Vol. 98, pp. 209-19, North Holland, Amsterdam (1989).

In one embodiment, floating gate 215 is formed from a GaN film grown in a 15 horizontal reactor operating at atmospheric pressure. Trimethyl gallium (TMG), trimethylaluminum (TMA), and ammonia ( $\text{NH}_3$ ) are used as source gases, and hydrogen ( $\text{H}_2$ ) is used as a carrier gas. The TMG, TMA, and  $\text{NH}_3$  are mixed just before the reactor, and the mixture is fed at high velocity (e.g., 110 cm/s) to a slanted substrate 230 through a delivery tube. The desired GaAlN composition  $v$  is obtained by controlling 20 the concentration ratio of TMG to TMA. In one embodiment, a 500 Å AlN buffer layer is obtained by growth at 600 degrees Celsius at a deposition rate of 100 Å/minute for approximately 5 minutes, then a epitaxial crystalline or polycrystalline layer of GaN is deposited at 1000 degrees Celsius.

In another embodiment plasma-enhanced molecular beam epitaxy (PEMBE) is 25 used to form a GaN or GaAlN floating gate 215, for example, by using electron cyclotron resonance (ECR) plasma during molecular beam epitaxy (MBE). The background pressure in the MBE chamber is typically less than  $10^{-10}$  torr. Ga flux (e.g., 99.9999% pure) is supplied by a conventional Knudsen effusion cell. The semiconductor substrates 230 are heated to a temperature of approximately 850 degrees

Celsius, and exposed to a nitrogen plasma (e.g., 35 Watt plasma power level) to clean the surface of the substrate 230 and form a thin AlN layer thereupon. The temperature is then lowered to approximately 550 degrees Celsius for growth of a thin (e.g., 300 Å) GaN buffer layer (e.g., using 20 Watt plasma power level for growth in a low active nitrogen overpressure environment). The temperature is then increased, such as to approximately 800 degrees Celsius, to form the remainder of the GaN or GaAlN film forming floating gate 225, such as at a deposition rate of approximately 0.22 microns/hour.

10

#### Conclusion

The present invention provides a DEAPROM cell. The memory cell has a gallium nitride (GaN) or gallium aluminum nitride (GaAlN) floating electrode, such as a floating gate electrode in a floating gate field-effect transistor. According to one aspect of the invention, a barrier energy between the floating electrode and the insulator is lower than the barrier energy between polysilicon and  $\text{SiO}_2$ , which is approximately 3.3 eV. The memory cell also provides large transconductance gain, which provides a more easily detected signal and reduces the required data storage capacitance value. According to another aspect of the invention, the shorter retention time of data charges on the floating electrode, resulting from the smaller barrier energy, is accommodated by refreshing the data charges on the floating electrode. By decreasing the data charge retention time and periodically refreshing the data, the write and erase operations can be several orders of magnitude faster. In this respect, the memory operates similar to a memory cell in DRAM, but avoids the process complexity, additional space needed, and other limitations of forming stacked or trench DRAM capacitors.

25

Although specific embodiments have been illustrated and described herein, those of ordinary skill in the art will appreciate that the above-described embodiments can be used in combination, and any arrangement which is calculated to achieve the same purpose may be substituted for the specific embodiment shown. This application is intended to cover any adaptations or variations of the present invention. Therefore, it is

manifestly intended that this invention be limited only by the claims and the equivalents thereof.